

**SYSTEM AND METHOD FOR MONITORING A
COMMUNICATION AND RETRIEVING INFORMATION RELEVANT TO
THE COMMUNICATION**

Field of the Invention

[001] This invention relates generally to the field of information retrieval, and more particularly to a system and method for monitoring a communication, and providing users with information retrieved from one or more data sources that may be relevant to the communication.

Background of the Invention

[002] The rise in popularity of the Internet, along with general advances in computer networking technology, has served to expand the possible means by which people may communicate and receive and disseminate information.

Electronic mail, newsgroup discussions, web-chats, and instant messaging, for example, illustrate means of communication that may be as common for some individuals as telephone calls or face-to-face meetings are for others.

[003] As a result, many individuals on a global level are being afforded increased opportunities to access information (or participate in virtual meetings, discussions, or other forums) that may cover a plurality of topics. Such opportunities, however, may be accompanied by an increased likelihood that individuals may receive information or be drawn into forums regarding topics for which they have little or no knowledge.

[004] In the business world, for example, it is not uncommon for executives or officials in large organizations to be copied on numerous electronic mail messages that may include unfamiliar acronyms, topics, or names relating to various topics, projects, or initiatives. To stay current and familiarize
5 themselves with industry buzzwords or terms of art, these executives may use a search engine to execute a search for relevant information across one or more networks.

[005] One drawback associated with many search engines, however, is that users are typically prompted to provide what they believe to be the most
10 appropriate keywords or search parameters. This may be disadvantageous in situations where users have little or no knowledge concerning a particular topic, or are unsure of how to structure the search.

[006] In addition, search engines, as well as other information retrieval systems and methods, may be of little value in situations where individuals need
15 real-time access to information. For example, it is not uncommon for business executives or other individuals to find themselves in a meeting or forum where the discussion has evolved into an area that they have little or no knowledge of. Moreover, it is not unusual for sales or customer representatives to find themselves in a client meeting where a question has arisen that they are unsure
20 of how to answer. In these and other situations, the probability of detouring a meeting to execute a search for information seems unlikely.

[007] These and other drawbacks exist.

Summary of the Invention

[008] The invention solving these and other problems relates to a system and method for monitoring a communication, and providing users with information retrieved from one or more data sources that may be relevant to the communication.

[009] According to an embodiment of the invention, a computer server, operating system, or other platform may host an information module. The information module may comprise a software program, an Internet web site, or other host site or application maintained by a service provider, or other entity.

The information module may, in one embodiment, comprise a contextual topic monitor (CTM) agent for monitoring a communication, and for providing users with information retrieved from one or more data sources that may be relevant to the communication. The CTM agent may further comprise a customization module, a monitoring module, a topic filter module, a search module, and a results module, each of which may implement the various features and functions of the invention as described herein.

[0010] In operation, one or more users may access the CTM agent through a remote client. The remote client may include any device associated with or present during a communication. According to an embodiment of the invention, a communication may include any real-time voice communication such as a telephone call or face-to-face conversation, regardless of the duration or number of participants involved. A communication may also include the transmission or receipt of a single text message, or the transmission or receipt of

a plurality of messages or postings comprising a real-time, text-based discussion, meeting, or other forum. Any other means by which information is exchanged may also be considered a communication.

[0011] According to an embodiment of the invention, the monitoring module of the CTM agent may receive a communication as input. Upon receiving a communication as input, the monitoring module may preferably forward the communication in real-time to the topic filter module.

[0012] According to an embodiment of the invention, the topic filter module may then analyze the communication to filter out and produce a running list of words that may define the context or key topics of the communication. A weighted averaging algorithm may be used to extract the relevant topic words in the communication. As the topic filter module continues to analyze the communication in real-time, it may update the running list so as to reflect the evolution of the topic(s) of the communication.

[0013] The search module may then utilize one or more of the topic words identified by topic filter module as search terms in a search executed against one or more data sources. The one or more data sources may comprise any combination of databases and other data storage or query formats, platforms, or resources. In this regard, a user of remote client does not have to formulate their own search terms, or even activate the search, while engaged in a communication.

[0014] After the search module has executed a search, the results module may format the results and transmit them to the remote client in real-time, according to one or more specified parameters. A results display may be presented to users for displaying the search results in any number of known
5 formats. For example, the results display may include links to documents or resources, such as hypertext links, hot links, or other known navigational tools that enable users to select, access, or display information. The results display may also include the full text of each document or resource identified during the search, or a brief synopsis of each identified search result. Other formats may
10 be presented.

[0015] For each result displayed in the results display, a user may have the option to save, print, or discard the result. Additional options may be presented to users depending on, for example, the particular remote client that a user has linked to the CTM agent. The results display may be updated to reflect the
15 evolution of the topic or topics of the communication. As described in detail below, the frequency of the update may be selected by users or established by the system.

[0016] One or more of the various features and functions of the modules comprising the CTM agent may be customized by a user, via a customization
20 module. A view or screen display associated with the customization module may be displayed when a user first accesses the CTM agent. This view may enable a user to select from various options. For example, a user may be permitted to select the types of communications that the monitoring module of

the CTM agent may monitor. Users may also be permitted to select from one or more of the various data sources available that the search module may execute a search against. In addition, users may specify whether search results are to be received via electronic mail, displayed on a designated intranet or Internet site,
5 or presented in a “pop-up” window, or other view. According to an embodiment of the invention, a user may designate one or more remote clients as recipients of the search results, and also specify the number of search results to be received at any one time.

[0017] One advantage provided by the invention is that the coupling of the
10 CTM agent to one or more data sources enables users of remote clients to have real-time access to resources that are relevant to the context of the communication to which they may currently be associated. This may empower the user with knowledge and expertise sources that match the evolution of the topic(s) of the communication.

15 *[0018]* Another advantage of the invention is that the topic filter module may analyze a communication to filter out and produce a running list of words that the search module may then utilize as search terms in a search executed against one or more data sources. Thus, a user is not required to activate a search, nor determine what search terms or parameters to use.

20 *[0019]* These and other objects, features, and advantages of the invention will be apparent through the detailed description of the preferred embodiments and the drawings attached hereto. It is also to be understood that both the

foregoing general description and the following detailed description are
exemplary and not restrictive of the scope of the invention.

Brief Description of the Drawings

5 [0020] FIG. 1 is a schematic diagram of a system for monitoring a
communication and providing users with information retrieved from one or
more data sources that may be relevant to the communication, according to an
embodiment of the invention.

10 [0021] FIG. 2 is a schematic diagram of an information module, according
to an embodiment of the invention.

[0022] FIG. 3 depicts a view illustrating an embodiment of a results
module, according to an embodiment of the invention.

[0023] FIG. 4 depicts a view illustrating an embodiment of a customization
module, according to an embodiment of the invention.

15 [0024] FIG. 5 illustrates a flowchart of processing according to the
invention, in one regard.

Detailed Description of the Present Invention

20 [0025] With reference to FIG. 1, a system 100 is provided, according to one
embodiment of the invention, which may comprise a server 150. Server 150
may be or include, for instance, a workstation running Microsoft Windows™
NT™, Microsoft Windows™ 2000, Unix, Linux, Xenix, IBM, AIX™, Hewlett-

Packard UX™, Novell Netware™, Sun Microsystems Solaris™, OS/2™, BeOS™, Mach, Apache, OpenStep™, or other operating system or platform.

[0026] An information module 200 may be stored on server 150, loaded on to server 150, or operated by server 150. Information module 200 may
5 comprise a software program, an Internet web site, an intranet site, or other host site or application maintained by a service provider, or other entity.

[0027] According to an embodiment of the invention, information module 200 may comprise a contextual topic monitor (CTM) agent for monitoring a communication, and providing users with information retrieved from one or
10 more data sources that may be relevant to the communication. As illustrated in FIG. 2, information module (or CTM agent) 200 may further comprise a customization module 210, a monitoring module 220, a topic filter module 240, a search module 260, and a results module 280, each of which may implement the various features and functions of the invention as described herein.
15 Additional modules may be provided. One or more of the modules comprising information module 200 may be combined. For some purposes, not all modules may be necessary.

[0028] As illustrated in FIG. 1, one or more data sources 160a – 160n may be operatively connected to server 150. Data sources 160a – 160n may
20 comprise any combination of databases and other data storage or query formats, platforms, or resources. Examples of databases may include the Oracle™ relational database sold commercially by Oracle Corporation, as well as Informix™ and DB2 (Database 2). Other databases may also be used.

Examples of additional data storage or query formats, platforms, or resources, may include a knowledge management (KM) repository, a search engine, OLAP (On Line Analytical Processing), SQL (Standard Language Query), a SAN (storage area network), and Microsoft Access™. Other data sources may also
5 be used, incorporated, or accessed into the invention.

[0029] In operation, one or more users may access information module 200 through an interface 115. According to an embodiment of the invention, interface 115 may comprise a graphical user interface (GUI) that may display various modules and functions available to a user via a remote client 110.

10 **[0030]** Remote client 110 may include any device associated with or present during the communication of information. For example, remote client 110 may comprise a personal computer, portable computer, PDA (personal digital assistant), workstation, dumb terminal, or other device that may be used to send, receive, or access any number of text-based messages. According to an
15 embodiment of the invention, these devices may also be configured with known hardware and/or software to enable voice recognition. A personal or portable computer having voice recognition capabilities that is functioning during a live business meeting or conference call, for example, may comprise a device associated with the communication of information. Additional devices,
20 including a web-enabled mobile phone, WAP device, or web-to-voice device are further examples of devices which may comprise remote client 110.

[0031] Remote client 110 may be networked directly to server 150, or connected over a network 130 to server 150 via a communications link 120.

[0032] Network 130 may include any one or more of, for instance, the Internet, an intranet, a PAN (Personal Area Network), a LAN (Local Area Network), a WAN (Wide Area Network), a SAN (Storage Area Network), or a MAN (Metropolitan Area Network).

5 **[0033]** Communications link 120 may include any one or more of, for instance, a copper telephone line, a Digital Subscriber Line (DSL) connection, a Digital Data Service (DDS) connection, an Ethernet connection, an Integrated Services Digital Network (ISDN) line, an analog modem connection, a cable modem connection, or a wireless connection.

10 **[0034]** Referring now to FIGS. 1 and 2, one or more individuals may participate in a communication associated with at least one remote client 110. Monitoring module 220 may monitor the communication by receiving either all (or a portion of) the communication as input. According to an embodiment of the invention, a communication may include any real-time voice
15 communication such as a telephone call or face-to-face conversation, regardless of the duration or number of participants involved. A communication may also include the transmission or receipt of a single text message (*e.g.*, an electronic mail message), or the transmission or receipt of a plurality of messages or postings comprising a real-time, text-based discussion, meeting, or other forum
20 (*e.g.*, a web chat). Any other means by which information is exchanged may also be considered a communication.

[0035] Upon receiving a communication as input, monitoring module 220 may preferably forward the communication in real-time to topic filter module 240.

[0036] According to an embodiment of the invention, topic filter module 240 may then analyze the communication to filter out and produce a running list of words that may define the context or key topics of the communication.

[0037] A weighted averaging algorithm may be used to determine one or more "likely active topics" (LAT) comprising a communication, based on extracted, relevant topic words from the communication. According to an embodiment of the invention, the weighted averaging algorithm may comprise an extraction process, a context-based refinement process, and an occurrence re-factoring process, each of which may be utilized to generate a LAT list for a particular communication, preferably in real time. Additional processes may be used. One or more processes may be combined and, in certain implementations, not all processes may be necessary.

[0038] According to an embodiment of the invention, an extraction process may be used to extract initial probable topics from the communication and return a raw topic vector. As described in greater detail above, a communication may comprise text-based communication (*e.g.*, electronic mail, web chat, text-messaging, etc.) or voice communication, among other communication formats.

[0039] With regard to text-based communications, the invention may incorporate existing topic-extraction algorithms for electronic text (*e.g.*, Lotus Discovery Server™) to generate a raw topic vector.

[0040] With regard to voice communications, the extraction process may
5 utilize or incorporate voice/speech recognition mechanisms (*e.g.*, a speech recognition engine) to resolve an audible sound input (*e.g.*, spoken word or words) against a dictionary or other word data source (not illustrated) to return a list of word matches in order of likelihood (*e.g.*, based on probabilities). For example, as an exemplary illustration, a voice communication may comprise a
10 spoken conversation between two individuals. One individual may utter the word “dog, ” and a speech recognition engine may generate a raw topic vector comprising a list of several potential matches for the word, with accompanying probabilities: fog?10%; log?18%; bog?13%; etc.

[0041] Currently, many voice/speech recognition systems focus on
15 recognizing one voice away from background noise, and attempt to refine the likelihood that a spoken word has been properly identified by the system based in part on lexical context of previous words or subsequent words using refinement post-processing techniques. Unfortunately, such systems that apply one-voice filtration and lexical refinements tend to yield unreliable results when
20 several voices are speaking, particularly if the system is trained against one voice. As such, according to an embodiment of the invention, one-voice refinement post-processing and/or filtering techniques may not be used. Rather,

a context-based refinement process (or multi-voice topic-averaging process)
may be performed on the raw topic vector.

[0042] In one implementation, the raw topic vector may be subjected to a
context-based refinement (or topic averaging) process that refines the raw topic
vector against the context of other near-in-time vectors, resulting in a refined
vector. According to one embodiment, and with further reference to the
exemplary “dog” illustration provided above, the topic averaging process may
compare recently generated raw topic vectors for a predetermined time interval
(e.g., for the previous two minutes) to strengthen the probability of the
likelihood that a “dog” or dogs is a topic of the conversation between the two
individuals.

[0043] With regard to text-based communications, the topic averaging
process may compare recently generated raw topic vectors for a predetermined
number of text-based characters (or other suitable metric) to strengthen the
probability that contextual topic word(s) are accurate. Accordingly, the
context-based refinement (or topic averaging) process may increase
probabilities for both voice communication and text-based communication by
15–20%.

[0044] Moreover, in various embodiments, the context-based refinement (or
topic averaging) process may further comprise additional refinement processes
or context filters such as, for instance, activity context, user context, taxonomy-
parent or synonym word look-up, involved-participant context, or topical
urgency context. Such additional refinement processes may enhance

probabilities for both voice communication and text-based communication by 50–100%. As an example, the aforementioned refinement processes or context filters may draw on user contexts such as mobile-phone location (when applicable and/or available), or profiles of participants listed for a given meeting on an e-calendar or other collaborative on-line meeting tool.

[0045] According to an embodiment of the invention, the refined vector that results from the context-based refinement process may then be subjected to an occurrence re-factoring process, thus resulting in a re-factored & refined vector. The re-factored & refined vector may be the subject of a search executed against one or more data sources as facilitated by the search module (described in greater detail below). It should be recognized that a high degree of certainty resulting from execution of the extraction and context-based refinement processes may indicate a high degree of certainty that a given topic is "active" for a given, current point-in-time topic vector. A final process performed on a refined vector may perform occurrence re-factoring. The invention may utilize tunable decay rate parameters and functions that can be adjusted against empirical tests for optimized default decay rates and curves, and then the occurrence re-factoring process may apply these curves to a current refined vector, deriving again from some predetermined time interval (or number of text-based characters) of prior vectors. A resulting re-factored & refined vector may include an additional relative priority value attached to each topic word. As an exemplary illustration, "dog,66%; cats,73%" may become "cats,66%,100%; cars,54%,50%; and dog,73%,30%". According to this

example, the “cats” topic was found more than any other topic within the decay bell-curve, so it is listed as the most active topic at 100% relative priority. The “cars” topic was injected into the current topic vector because it exists in older topic vectors at accuracy rating 54% and its importance may not decay completely to 0%, for example, for another 3 minutes. If “cars” as a topic does not appear again, it may appear in subsequent vectors with lower & lower priority until 0% priority is reached, at which point it may be dropped. If “cars” appears occasionally, then the decay rate may slow or may even reverse. Additionally, if the occurrence rate becomes high enough for “cars,” then the relative priority for this topic may climb. Finally, in this example, “dogs” remained in the vector but its frequency of prior occurrence is so low that its priority drops. As such, in an exemplary illustration of a live voice conversation, or multi-way text chat session between numerous participants may yield the following results: (pass1) → "dogs,20%; cats,18%;" (pass2) → "dogs,66%; cats,73%;" (pass3) → "cats,66%,100%; cars,54%,50%; dog,73%,30%." This kind of vector, recalculated at a given time interval (*e.g.*, every second), may guide searches against one or more data sources as described below.

[0046] According to an embodiment of the invention, search module 260 may utilize one or more of the topic words identified by topic filter module 220 as search terms in a search executed against one or more of data sources 160a – 160n. In this regard, a user of remote client 110 does not have to formulate their own search terms, or even activate the search, while engaged in a

communication. By coupling information module (CTM agent) 200 to one or more data sources 160a – 160n, a user of remote client 110 may be presented with access to knowledge reports of experts, various documents, information, or other resources relevant to the context of the communication to which they may currently be associated. This may empower the user with knowledge and expertise sources that match the evolution of the topic(s) of the communication.

[0047] After search module 260 has executed a search, results module 280 may format the results and transmit them to remote client 110 in real-time, according to one or more specified parameters. FIG. 3 depicts a view 300 illustrating an embodiment of results module 280 according to an embodiment of the invention. View 300 may comprise a topic word display 320 and a results display 340. Additional displays may be presented in view 300. Topic word display 320 may display one or more of the topic words generated by topic filter module 240 that were used as search words by search module 260. Results display 340 may display any number of results in any number of known formats. For example, results display 340 may include links to documents or resources, such as hypertext links, hot links, or other known navigational tools that enable users to select, access, or display information. Results display 340 may also include the full text of each document or resource identified during the search, or a brief synopsis of each identified search result. Other formats may be presented.

[0048] For each result displayed in results display 340, a user may have the option to save, print, or discard the result by selecting one or more of selection

portions 360a, 360b, and 360c respectively. Additional options may be presented to users through additional selection portions. The selection portions available to users, as well as the means used to select them, may vary dependent upon the particular remote client 110 used to access information module 200 and receive results. For example, in an embodiment where remote client 110 comprises a personal or portable computer, a user may select one or more of the selection portions (360a, 360b, 360c) using a keyboard, mouse, or other known input device. In an embodiment where remote client 110 comprises a web-enabled mobile phone, however, results display 340 may prompt a user to press a key on the numeric keypad of the phone to save or discard a result.

[0049] Both the topic word display 320 and results display 340 may be updated to reflect the evolution of the topic or topics of the communication. As described in detail below, the frequency of the update may be selected by users or established by the system. If a user has selected to save a result in results display 340, the result may continue to be displayed even after the display has been updated with new search results.

[0050] One or more of the various features and functions of the modules comprising information module 200 may be customized by a user, via customization module 210. FIG. 4 depicts a view 400 illustrating an embodiment of customization module 210, according to an embodiment of the invention.

[0051] View 400 may be displayed when a user first accesses information module 200. View 400 may, according to one implementation, include a logo

box 410 for displaying text, a logo, or any other icon, symbol, or graphic. Logo box 410 may serve an aesthetic function, or may display an advertisement, or a trademark (or graphic) identifying the service provider or entity maintaining information module 200. One or more logo boxes 410 may be present. View
5 400 may further comprise a first data field 420, second data field 430, and a third data field 440, each of which may include various selection portions that enable users to specify preferences. Other data fields may be present. In addition, the selection portions available to users in the various data fields may vary dependent upon the particular remote client 110 used to access information
10 module 200.

[0052] According to an embodiment of the invention, first data field 420 may allow a user to select the types of communications that monitoring module 220 may monitor. A user may, for example, choose to monitor text-based communications, voice communications, or other types of communications
15 associated with a remote client 110. A user may also wish to monitor all types of communication associated with a remote client 110.

[0053] According to an embodiment of the invention, second data field 430 may allow a user to select from one or more of the various data sources 160a-160n available that search module 260 may execute search against. A user may
20 also select the frequency with which topic filter module 240 updates the list of topic words and executes a search. The update frequency may comprise any interval of time.

[0054] Third data field 440 may enable a user to select from a number of formats for which to display search results associated with results module 280, as previously described. For example, a user may wish to receive results via electronic mail, or have results displayed on a designated intranet or Internet site. In addition, results may also be presented in a “pop-up” window or other view. Other display options may be provided.

[0055] According to an embodiment of the invention, a user may also specify one or more remote clients 110 (other than that associated with the communication) to receive results. As an illustrative example, assume that a business executive in a meeting is monitoring the meeting discussion via a voice-enabled laptop computer. The business executive may wish to have the search results displayed only on the laptop screen. In this regard, as the results on the laptop screen are quietly updated, the business executive may visually scan the screen without breaking the discussion flow of the meeting. The business executive may, however, wish to have the results presented on a large monitor for each of the meeting participants to see.

[0056] According to one implementation, users may also specify the number of search results to be received at any one time. Once a user's options have been selected, a user may submit them by selecting a submit selection portion 450. Selections may also be voided by selecting a cancel selection portion 460.

[0057] FIG. 5 illustrates a flowchart of processing according to the invention, in one regard. Processing may begin at step 500. In an operation

504, a monitoring module associated with an information module may monitor a communication associated with at least one remote client by receiving either all (or a portion of) the communication as input. The information module may comprise a software program, an Internet web site, an intranet site, or other host site or application maintained by a service provider, or other entity. In addition to the monitoring module, the information module may further comprise a customization module, topic filter module, search module, and results module, each of which may implement the various features and functions of the invention. Additional modules may be provided.

10 *[0058]* According to an embodiment of the invention, the communication received may include any real-time voice communication such as a telephone call or face-to-face conversation, regardless of the duration or number of participants involved. The communication may also include the transmission or receipt of a single text message, or the transmission or receipt of a plurality of messages or postings comprising a real-time, text-based discussion, meeting, or other forum. Any other means by which information is exchanged may also be considered a communication. Upon receiving a communication, the monitoring module may preferably forward the communication to a topic filter module for further processing. In an operation 508, the topic filter module may analyze the communication to filter out and produce a running list of words that may define the context or key topics of the communication. A weighted averaging algorithm may be used to extract the relevant topic words in the communication. As the topic filter module continues to receive the

15

20

communication from the monitoring module in real-time, it may update the running list so as to reflect the evolution of the topic(s) of the communication.

5 *[0059]* Once one or more topic words have been identified by the topic filter module, the search module may, in an operation 512, utilize one or more of the topic words as search terms in a search executed against one or more data sources. Available data sources may comprise any combination of databases and other data storage or query formats, platforms, or resources. In this regard, a user of a remote client does not have to formulate their own search terms, or even activate the search, while engaged in a communication. By coupling the information module to one or more data sources, a user of a remote client may be presented with access to knowledge reports of experts, various documents, information, or other resources relevant to the context of the communication to which they may currently be associated. This may empower the user with knowledge and expertise sources that match the evolution of the topic(s) of the communication.

10

15

20 *[0060]* In an operation 516, after the search module has executed a search, the results module may format the results and transmit them to remote client in real-time, according to one or more specified parameters. A results display, for example, may be presented to users. The result display may include links to documents or resources, such as hypertext links, hot links, or other known navigational tools that enable users to select, access, or display information. The results display may also include the full text of each document or resource

identified during the search, or a brief synopsis of each identified search result.

Other formats may be presented.

[0061] According to an embodiment of the invention, an inquiry may be made in an operation 520 to determine whether the communication is complete, or is continuing to evolve. If a communication is complete, processing may end at stop 524. Otherwise, processing may be re-entered at step 504.

[0062] Other embodiments, uses and advantages of the invention will be apparent to those skilled in the art from consideration of the specification and practice of the invention disclosed herein. The specification should be considered exemplary only, and the scope of the invention is accordingly intended to be limited only by the following claims.